


Weekly Report (2019.7.8-2019.7.14)

DONE

1. VAST论文修改final版本

1. 周三收到论文录用的邮件，然后开始修改按照要求修改论文的 final 版本，添加作者及脚注+添加致谢章节（基金号）+去除论文中的红色标记+删减篇幅+添加参考文献。另外，还在视频添加作者列表。

Decision on submission 1065 of VIS 2019weiyating@zju.edu.cn



Remco Chang, Daniel Keim, and Ross Maciejewski 2019-07-09 12:17
发至 我
(此邮件由vis19a+log.5d241515b7371870b4b5f987@precisionconference.com代发)

🔍 ⏪ ⏩ ⋮
[详情](#)

Dear Yating Wei,

Congratulations! We are delighted to inform you that your paper

1065 - Evaluating Perceptual Bias During Geometric Scaling of Scatterplots

was definitively accepted for presentation at IEEE VIS 2019 VAST Papers, to be held October 20-25 in Vancouver, Canada. Your paper will be published as a special issue of IEEE Transactions on Visualization and Computer Graphics (TVCG).

2. 联邦学习可视化

1. 参与林昊的联邦学习可视化的讨论，通过不断发问，一起梳理清楚系统设计逻辑。修订设计文档，并讨论分工。考虑到联邦学习可视化的开发需求要比联邦可视化的复杂，我让舒悦帮忙开发，朱融晨小学妹做联邦可视化的前端开发。

3. 联邦可视化

1. 周三的时候，与智勇讨论他近期的调研结果（我们不一定用联邦学习的中心平均算法，也许可以根据我们的具体需求去做，而且在通信开销与模型准确性之间有一个counterbalance，后续在做的过程中要有一个平衡）这也许是可以考虑的一个细节点。
2. 周三完成final版本的修订，周四也加入调研，并讨论接下来的工作。调研hanwei shen老师的工作

在原位处理部分，作者提出使用分布数据来近似表示原始数据，从而实现对数据的约减，也同时保留重要信息。具体来说，原始模拟的空间域被划分为若干个小块 (block)，对于每个时间步里的每个块，用分布来表示原始数据。这里，作者选用高斯混合模型 (Gaussian Mixture Model, GMM) 来表示分布数据，即用若干个不同权重的高斯核来拟合数据的分布。

$$p(x) = \sum_{j=1}^K \omega_j * \mathcal{N}(x|\mu_j, \sigma_j)$$

图3：高斯混合模型

作者之所以选择高斯混合模型，主要是基于其在存储上的高效率。一些无参数的模型，例如核密度估计，需要高昂的计算的存储代价；直方图能很快的计算，但存储代价高；有参数模型，如使用单个高斯核，则可能导致不精确。

但是，GMM有点老，我们采纳陈老师的意见，继续用我们的神经网络模型来训练拟合模型。

3. 接下来的工作（日程表见短期计划）a)想法：随着人们的隐私保护意识的日益加强，以及公司间利益的纠葛，很多时候，我们不能简单的将数据整合起来进行数据分析。受到联邦学习启发，我们提出一套解决方案，能够在保障数据隐私的前提下，对分布在不同机构的数据进行可视分析。b) 实现：智勇负责后端开发，根据不同任务，训练出几个神经网络模型。朱融晨负责前端开发，将后端吐出的数据绘制成热力图、柱状图这些图表。周舒悦从暑期开始从 联邦学习可视化项目 回归到联邦可视化项目，做前端，并帮助我修订论文写作。c) 实验数据：滴滴开源数据，现在手上有西安出租车轨迹数据，另外我们还申请另外两份数据：有起始点的出行数据以及POI数据，不过这些数据不是同一地点的，不能叠加起来分析。d) 写作：我来写，之前写短文的那些文字可以拿过来用。另外，想请梅博给我的写作把把关，在结构上进行指导。

4. 其他

- 1. 完成 MOOC 弹题制作
- 2. 做 dataV 的 user study
- 3. 带领暑期实习小学妹朱融晨办理宿舍入住

小结

工作日工作时长9+11+12+8+3，周末5,总时长约48h。周四傍晚，老公去做了一台小手术，手术后人动不了，我吃完晚饭就赶过去陪护了，周五请了一天假。这几天工作有些耽搁了。

PLAN

短期计划（一周）

- 1. 组会报告博客撰写
- 2. 第七章书籍修订
- 3. VIS开会日程表
- 4. 联邦可视化投稿日程安排（后序需要再细化）

时间	任务	其他
0715-0717	确定研究任务，然后设计系统界面	
0718-0720	introduction & 系统实现	20号出发去成都
0721-0723	related work	这三天都在成都，完成的工作会很少

0724-0726	related work & 系统实现	
0727-0729	框架设计 & 系统实现	27、28参加同学婚礼
0730-0801	框架设计、实现 & 系统实现	
0802-0804	框架实现 & 系统实现	
0805-0807	case study	
0808-0810	case study	
0811-0813	discussion、conclusion	
0814-0816	全篇改	
0817-0819	全篇改	

中期计划

1. 专利两篇（1. 联邦可视化框架：舒悦 2. ? ）
2. 完成 Visevo 投稿项目
3. 准备金融可视化的综述

长期计划

1. 学习更多机器学习、数据挖掘相关的算法。
2. 在项目中锻炼自己的思考能力与代码能力。